AD-A209 984

1b RESTRICTIVE MARKINGS

3 DISTRIBUTION / AVAILABILITY OF REPORT
Approved for public release; distribution unlimited.

4 PERFORMING ORGANIZATION REPORT NUMBER(S)

5 MONITORING ORGANIZATION REPORT NUMBER(S)
AFOSR-TR- 89-0975

6a NAME OF PERFORMING ORGANIZATION
New York University

6b OFFICE SYMBOL (If applicable)

7a NAME OF MONITORING ORGANIZATION
Air Force Office of Scientific Research/NL

6c ADDRESS (City, State, and ZIP Code)
550 1st Avenue
New York, NY 10016

7b ADDRESS (City, State, and ZIP Code)
Building 410
Bolling AFB, DC 20332-6448

8a NAME OF FUNDING / SPONSORING ORGANIZATION
AFOSR

8b OFFICE SYMBOL (If applicable)
NL

9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
AFOSR-85-0341

8c ADDRESS (City, State, and ZIP Code)
Building 410
Bolling AFB, DC 20332-6448

10 SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION N
61102F	2313	A5	

11 TITLE (Include Security Classification)
Shape Description with a Space Variant Sensor: Algorithms for Scan-Path, Fusion and Convergence Over Multiple Scans

12 PERSONAL AUTHOR(S)
Eric Schwartz

13a TYPE OF REPORT	13b TIME COVERED FROM TO	14 DATE OF REPORT (Year, Month, Day)	15 PAGE COUNT
Reprint			22

16 SUPPLEMENTARY NOTATION

17 COSATI CODES		
FIELD	GROUP	SUB-GROUP

18 SUBJECT TERMS (Continue on reverse if necessary and identify by block number)

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

20 DISTRIBUTION / AVAILABILITY OF ABSTRACT
☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS

21 ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED

22a NAME OF RESPONSIBLE INDIVIDUAL
John F. Tangney

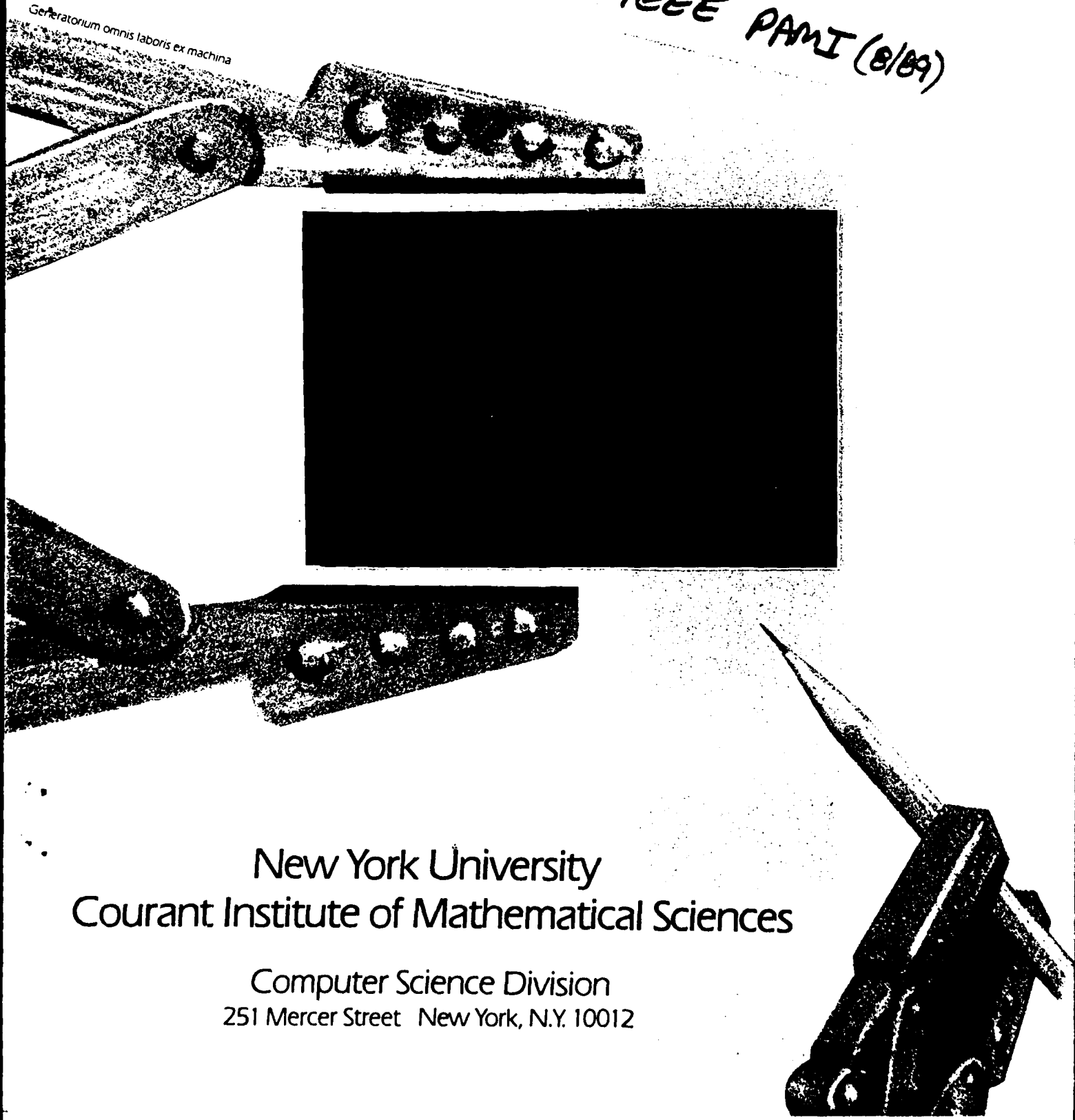
22b TELEPHONE (Include Area Code)
(202) 767-5021

22c OFFICE SYMBOL
NL

DD FORM 1473, 84 MAR 83 APR edition may be used until exhausted.
All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

Robotics Research Technical Report



New York University
Courant Institute of Mathematical Sciences

Computer Science Division
251 Mercer Street New York, N.Y. 10012

Shape Description With a Space Variant Sensor: Algorithms for Scan-Path, Fusion and Convergence Over Multiple Scans

by

Yehezkel Yeshurun †
Eric L. Schwartz ‡†

Technical Report No. 295
Robotics Report No. 109
April, 1987

‡ New York University
Dept. of Computer Science
Courant Institute of Mathematical Sciences
251 Mercer Street
New York, New York 10012

† Computational Neuroscience Laboratories
Dept. of Psychiatry
NYU Medical Center
550 First Ave
New York, N.Y. 10016

Work on this paper has been supported by the System Development Foundation, AFOSR #85-0341, and the Nathan S. Kline Research Institute.

"@(#)version.1 1.6 4/15/87 NYU Comp. Neuro Science Lab"

SHAPE DESCRIPTION WITH A SPACE VARIANT SENSOR: ALGORITHMS FOR SCAN-PATH, FUSION AND CONVERGENCE OVER MULTIPLE SCANS

Yehezkel Yeshurun

Eric L. Schwartz

Department of Computer Science
Courant Institute of Mathematical Sciences
715 Broadway
New York, N.Y. 10003

Computational Neuroscience Laboratories
New York University School of Medicine
Department of Psychiatry
550 First Avenue
New York, N.Y. 10016

ABSTRACT

One of the ways by which early human vision is sharply distinguished from machine vision is by the fact that the human visual representation is strongly space variant and that the human system builds up a representation of a scene through multiple fixations during scanning.

In this paper, we discuss three algorithms related to the "blending" of a single scene from multiple frames acquired from a space variant sensor.

1.) Given a series of space-variant contour based scenes, with different "fixation points", we show how to fuse these into a single, multi-scan view, which incorporates the information present in the individual scans.

2.) We demonstrate an (attentional) algorithm which recursively examines the current knowledge of the scene, in order to best choose the next fixation point, based on focusing attention in regions of maximum boundary curvature.

3.) We discuss a simple metric for evaluating "convergence" over scan-path. This may be used to quantify the performance of (2) above, i.e. to compare the performance of various "attentional" algorithms.

Finally, we discuss this work in the light of both machine and biological vision.

SHAPE DESCRIPTION WITH A SPACE VARIANT SENSOR: ALGORITHMS FOR SCAN-PATH, FUSION AND CONVERGENCE OVER MULTIPLE SCANS

Yehezkel Yeshurun

Eric L. Schwartz

Department of Computer Science
Courant Institute of Mathematical Sciences
715 Broadway
New York, N.Y. 10003

Computational Neuroscience Laboratories
New York University School of Medicine
Department of Psychiatry
550 First Avenue
New York, N.Y. 10016

Introduction

When we view a scene, we have the subjective impression that what we see is stable and constant, both in position and resolution. However, this impression is far from correct. If we try to read a newspaper that is slightly off-center (see Fig.1), we become aware that the very high resolution provided in the region of our fixation (foveal projection) falls off rapidly toward the edges of our field of vision. The fact that the human visual representation is strongly space variant implies that the human system builds up a representation of a scene through multiple fixations during scanning.

The space variant nature of the human visual system is well understood, at least to the level of primary visual cortex. The threshold for visual acuity, stereo-acuity, motion, and other psychophysical quantities scale at least roughly as the inverse of dis-

supported by AFOSR #F 85-0235, System Development Foundation and the Nathan S. Kline Psychiatric Research Center

tance from the fovea. There is general consensus[1, 2, 3] that the spatial representation of the visual field [1] , at the level of the primary visual cortex, is approximated by a complex logarithmic mapping[4]. Figure 1 and Figure 6 of this paper show natural scenes processed by this form of mapping function. We are thus in a position to provide realistic estimates of the nature of a specific space variant imaging system: that of the human.

In the present paper, we discuss three algorithms related to the "blending" of a single scene from multiple frames acquired from a space variant sensor. We used contour based scenes, rather than gray scale scenes, in order to focus attention on the problem space variance, as opposed to segmentation. The following generic problems are raised by considering a a space variant system:

1.) Given a series of space-variant contour based scenes, with different "fixation points", how might one fuse these into a single, multi-scan view, which incorporates the information present in the individual scans?

2.) How might one choose successive fixations points, in order to rapidly gather shape dependent data? Is there a simple attentional algorithm for contour based scenes?

3.) How could one quantify the rate of convergence of such a system, as a function of the number of scans? What is the rate of convergence suggested by such a metric? [2]

In the present work, we do not address the classical issues of how the system (human or machine) is to obtain knowledge of its motor state (see 5). Our intention here is to discuss the image processing problem of blending together multiple scans, obtained from a strongly space variant sensor, and the problem of choosing a "scan

---

[1] In this paper, we do not discuss the detailed spatial architecture of primary visual cortex, which would include details such as ocular dominance columns, orientation columns, etc. We are only concerned here with the first order topographic structure of the human visual system, as a model for space variant machine vision systems.

[2] In addition to these purely computational issues, the human system has also needed to: 1.) evolve systems of accurate motor control, 2.) provide information to the organism about the current motor state (i.e. direction of gaze). This aspect of the problem has been much discussed under the terms proprioceptive perception, efference copy, corallary discharge, etc.[5].

path" which provides optimal information about the scene.

Another interpretation of the work described here might be made, entirely within the context of machine vision. Assuming that a space variant sensor similar to a human retina were available, it would be necessary to consider some of the issues discussed in the present paper: how should one choose a series of fixation points for such a sensor, how would one blend the successive frames, and how could one place a metric on the quality of this scanning process?

The space-variant image and boundary-angle function

We define the resolution at the point v of an image as the function $R_p(v)$, where p is the spatial location of a fixation point and R is a monotonic non-increasing function of $|v-p|$. This is to say that R is proportional to the reciprocal of the minimal distinguishable distance (i.e visual acuity). In the current context the exact specification of R is not crucial; any R having the mentioned attribute can be used. The following discussion uses a function of the form $\frac{c}{|v-p|}$, for $v \neq p$, where c is a constant.

This definition might be applied to any gray-scale image (see Fig. 1). In the current application we consider only contour based images. This situation can arise either naturally, when a scene is two-dimensional and consists only of contours, or artificially, after an edge-detection mechanism has been applied to an image of a complex three-dimensional scene (segmentation).

Boundary contour descriptor

In applications in which a one-dimensional representation of contours is desired, it is customary to use the boundary-angle function $\theta(l)$, which is the angle of the tangent to the contour, as a function of the arc-length unit l. In the current application, since we have discrete points connected by line segments (i.e polygons) , we use the representation $\Theta(l)$, which is the difference between two consecutive angles of the polygon. This one-dimensional representation of contours is most useful in shape-

recognition tasks, where it is further processed by a Fourier transform to yield the Fourier descriptors (FDs) of the contour [6]. There are also some indications that the FD of a shape might be useful as a shape descriptor in physiological studies of the primate visual system[7].

We apply, spatial-variant resolution to both the image of the contour in the x/y plane and to the boundary-angle $\Theta(l)$ representation of it, as explained below (see also Figure 2).

1) The original contour is represented by line segments between the points $\{U_i, i=1,k\}$. We assume that the distance between these points represents the highest possible resolution of the "viewer."

2) A new contour is defined by a fixation point: Given a fixation point p, and a contour point U_i, the value of $R_p(U_i)$ determines the next point U_j. Thus, starting at U_0, this procedure yields a contour whose points are a subset of the original points.

3) The boundary angle of the new contour, $\Theta_p(U_i)$, $i\varepsilon\{1,k\}$, is obtained. To allow reconstruction of the original image, we also keep the resolution value $R_p(U_i)$ for each U_i.

In the x/y plane, variable resolution produces a detailed image near the fixation point and a "blurred" image away from the fixation point. In the boundary-angle representation, the neighborhood of the fixation point is properly described, while other areas retain only smoothed, low-frequency details. The parameters used in this work yield a ratio of 1:10 between the full resolution image and a single space-variant view, which is in good agreement with the functional form of human visual acuity[3].

[3] One recent estimate of primate magnification factor[1] suggests that there is a 10:1 decrease in spatial resolution of a stimulus between the fovea and five degrees of eccentricity. This is a reasonable "viewing aperture" for shape perception. Note that a 10:1 (linear) change corresponds to a 100:1 area change, and that this area change is a more relevant index of "data compression".

Blending boundary-angle functions and images

For a given fixation point, there exists a corresponding representation of the original contour. Several fixation points $\{p=p_1 \cdots p_n\}$ produce different representations of the same contour. This situation is shown in Figure 3, in which images are viewed from several different points. Although the boundary-angle function $\Theta_p(U_i)$ is quite detailed near the corresponding fixation point, it just roughly approximates the original boundary-angle function in all the other areas.

Because resolution depends only on the distance between a given point and the fixation point, and because the most detailed boundary functions (or images) are obtained for high-resolution areas, an appropriate blending scheme should use the "best" of each view. The only information the blending scheme needs is the resolution associated with each point in the subcontour, which is kept when the subcontour is calculated. Thus, the reconstructed boundary-angle function is

$$\Theta^*(U_l) = \Theta_f(U_l)$$

such that

$$R_f(U_l) = \max_{p=p_1 \cdots p_n} \{R_p(U_l)\}.$$

The reconstructed function $\Theta^*(l)$ is an approximation to the original $\Theta(l)$. This approximation depends on the number of fixation points and their location. A more elaborate blending scheme might also depend on the "scanpath" or sequence of fixation points humans select when viewing a given scene[8].

Choice of scan path: an "attentional" algorithm

Although early vision and artificial intelligence (late vision?) have received a great deal of attention recently, a great intermediate area exists which has received little study in this context, and that is the subject of ''attention'' itself. A single scan provides partial information about a scene. Assuming that a unified representation of

the scene can be extracted from successive scan, we must address the problem of locating the fixation points,in such a way as to provide maximal information to the imaging system. This represents an ill defined problem, as difficult issues relating to context and goal direction are implied by it. However, little advantage can be gained from a space variant system without providing an attentional algorithm. In the following, we will discuss a simple candidate for attentional choice of successive fixation points.

In psychophysical contexts, the nature of visual scanning has been extensively explored (e.g., 9). In general, fixation points tend to cluster around sharp edges, ends of lines, and locations where some "unpredictable" change takes place. Although most existing research considers only the question of the location of the fixation points, some of the literature does pay attention to the temporal ordering of these points, which is termed the "scanpath"[8].

In our case, the scene consists of contours. The curvature of the contours is very likely to be a prime fixation-point "attractor", since large curvature represents rapid rate of change of boundary orientation. We can represent the curvature in terms of a boundary-angle function, indicating areas of high curvature by corresponding peaks in the function. A simple form of attentional algorithm, then, consists of the following steps:

1) Chose (randomly, or by any method) an initial fixation point.

2) Calculate the boundary-angle function according to the current fixation point.

3) Select the next fixation point according to the maximum of the boundary-angle function $\Theta_p(U_i)$.

4) Keep the boundary angle function and the corresponding resolution values. Keep a reference point in the current fixation, that will be associated with a point in the next fixation.

5) Blend the views and the boundary angle functions to yield a single view/function.

6) Go to step 2, until "convergence" (see below).

Such a procedure is shown in Figure 4. The fixation points in this figure seem plausible in comparison with the points that one would likely select without using the algorithm. However, the algorithm has one drawback. In cases where several high values of the boundary-angle function cluster together, the algorithm picks several fixation points at almost the same place. Because the scans obtained from adjacent fixation points do not differ much, and because the foveal area can cover several points of high curvature, this clustering of points is redundant.

In order to remove the redundancy, we modify the algorithm (in step 3) by considering $\Theta(U_i)W(U_i)$ instead of $\Theta(U_i)$. The weight function $W(U_i)$ can be used to enhance (or mask) selected features. If W is chosen such that it equals 1 everywhere except for a neighborhood of the fixation point where it vanishes, the redundancy problem is solved. In other words, after a fixation point is selected, the relevant foveal area (i.e., the area immediately surrounding the fixation point, where the high resolution still holds) is not counted when the algorithm searches for the next-higher value. Figure 4b shows the results of this approach.

One might also select W to be $\frac{1}{R}$, thus emphasizing "remote" features rather than "close" ones. Finally, W might contain some random fluctuations, in order to avoid the possibility of being "trapped" between two features.

The algorithm needs a reference point that is shared between each two succesive fixations: this is necessary when the views, or the boundary angle functions, are "tailored" together.

Convergence and norms

Because our figures consist of simple contour drawings, it is easy to define a norm that compares composite space-variant scenes after n scans with the original high-resolution scene. A reasonable choice for this norm is a least-squares measure of the two boundary-angle functions. Thus, let Δ_n represent the difference between the full-resolution scene and the composite scene after the incorporation of the n^{th} fixation point : $\Delta_n = |U - C_n|$.

Using this norm, it is possible to define the convergence rate as a function of the scanpath. Thus, for a sequence of fixation points $p_1, p_2, \ldots p_n$, we define the rate of convergence for the scan path at point n, as $\Delta_n - \Delta_{n-1}$. This method is suitable for the purpose of the algorithms evaluation or for calibration, when we have access to the full resolution contour. However, in a "real-time" situation (i.e in robotic vision), the full resolution image is not necessarily available. Thus, we can define Δ_n as $|C_n - C_{n-1}|$, and base the "convergence" decision on it (see Fig. 5). If one thinks of n as a time variable then this measure indicates the ''rate'' of error-reduction.

Thus, one algorithm for adding scanpaths might be based on the addition of a new point which, among all the possible fixation points, maximizes the above ''rate'' of convergence. Conversely, the addition of new points becomes unecessary when no points can be found that significantly improve the rate of convergence. The algorithm we propose rapdily converges: it is monotonic, in the sense that only "better" resolution points are introduced, and it is bounded by the original set of points which constitutes the object. Figure 5 shows an example of an aircraft silhouette which is scanned by this algorithm, with a plot of convergence based on the latter method described above. It is clear that there is rapid convergence to an accurate representation of the boundary of the figure. It is interesting to note that [8] report that humans typically view scenes with perhaps 3 - 8 scans; our algorithm also converges quite rapidly, in this case in which parameters of space variance derived from human vision have been

used.

In more general cases, however, the choice of a norm is likely to be quite difficult. In the general case, both the attentional algorithm and the norm used to evaluate its success would likely be dependent on past experience, the goal-directed state of the imaging entity, and the full context of the current task. In lieu of engaging in this full-blown algorithmic study of visual attention, we propose that the simple curvature based norm and scanning algorithm outlined above provides an initial step in the direction of understanding visual attention, and is one which is optimal in those situations in which a value-neutral estimate of boundary curvature is the desired information.

Implication of space variant image processing to gray-level images.

Though we address mainly contour-based images in this work, it might be of interest to point out its application to gray-level images, especially from the aspect of "data compression".

The human visual field subtends roughly 100 x 100 degrees[10] , with a maximum resolution of about 1 minute of arc (foveal). Using a space invariant sensor (e.g. conventional CCD camera), one would have to resolve 6000x6000 samples (1 minute of arc x 100 degrees in each direction). In order to achieve this performance, one would have to sample at 2-3 times this resolution, in each dimension. An image of 16000x16000 would provide this performance, but would extend close to the gigapixel range in size.

We have experimentally demonstrated this estimate by digitizing[4] a conventional eye-chart, at a distance of 20 feet, using a wide angle (fisheye) lens, which recorded from about 80 degrees of field. Figure 6 shows the "full scene", and a highly

[4] We used a conventional NTSC frame grabber, at 480x525 resolution, together with a polar coordinate mosaic technique[11] to produce this simulation.

magnified detail of the eye-chart, at the center. We continued to magnify the scene, until the 20/20 line of the eye-chart was visible (indicating a resolution of about 1 minute/arc). We calculate that this occured at an effective sampling resolution of 16,0000x16,0000 pixels.

Although both of the previous estimates are ad-hoc, they agree well enough to suggest that the effective resolution of a single scan of the human system is equivalent, were it recorded by a space invariant system, to a 1/4 giga-pixel image. Now, this estimate of 1/4 giga-pixel is based on the use of a constant resolution system, which extended over 100x100 degrees, at full visual acuity. In fact, we simulated the logarithmic structure of the human visual system, and our simulated image occupied only about 16000 pixels (see figure 6). Naturally, we only obtained high resolution over a small "foveal" representation with this simulation; in order to use this approach effectively, multiple scans would need to be performed. However, with a relative data compression of about 16,000 : 1 , we can afford to perform the scanning process over a number of fixation points. Even 16 sucessive fixations would yield an effective 1000:1 data compression relative to a constant resolution system, provided that one obtained a satisfactory representation of the image regions of interest.

Summary

Space variant imaging has been little explored in the context of machine vision, but is a major area of interest in the context of biological vision. Space variant imaging provides a number of advantages, and difficulties, with respect to conventional space invariant systems. One advantage is that very large fields of view can be covered, and very high resolution can also be provided.This leads to a form of image data compression which can be extremely large. However, a number of algorithmic difficulties are introduced by considering strongly space variant systems. Attentional algorithms are required to make effective use of the small high resolution "fovea", while other algorithms are required to "fuse" successive space variant scans.

In the present paper, we have provided preliminary solutions to each of these issues. Using our algorithms, we obtain satisfactory convergence, for reasonable parameters of space variance derived from human vision, over a small number of scans (perhaps 3-5 scans).

The possibility that space variant sensors (e.g. CCD's) may become available for application in machine and robotic vision should provide some motivation to begin studying the issues which such a sensor would provide. Perhaps the possibility that some of the high performance of the human visual system derives from its use of a space variant architecture may provide some impetus to develop such a sensor.

Figure Captions

Figure 1. Figure 1A simulates six successive scans of a newspaper, using a cortical map function derived from primate data[6] ,a reading distance of about twenty centimeters, and about 1.5 degrees of visual field on each side of the fixation point. Each of the small "bow ties" represents the cortical "image" of a section of newspaper print. Thus, the first frame is fixated on the letter "o" in the word "roaches". There are two "bow ties" representing the left and right visual fields. The newspaper is then scanned, and the corresponding cortical "images" are presented in the figure. Note the strong space variance, even for the central few degrees of visual field.

Figure 1B shows these six scans projected back to the visual field, and "fused" into a single scene[13]. The region of text scanned, which read " roaches don't die..", and too some extent the lines above and below this line, are seen clearly, but there is a rapid loss of detail in the text regions which are not close to the scanned text. Figure 6 of this paper shows a wide angle simulation of the human visual field and cortical image.

Figure 2. A: Images (left) and their boundary-angle functions (right). Top: the original contour (black silhouette) and its boundary-angle function. Bottom: the image as it is "viewed" from the fixation point (indicated by a star), with space-variant resolution. The tail of the airplane, being fairly far from the fixation point, is described very roughly. Therefore, the boundary-angle function bears only a rough resemblance to the original function.

B: A scene consisting of several planes sillhouetes (a), as it is "received" from different fixation points (b-d). The fixation points are depicted by an asterix. The original airplane silhouette consists of 243 points, and the space-variant silhouettes average 5 points (for the less detailed ones) to 40 points (for the highly detailed).

Figure 3. A: View of a triangle from three fixation points. The contour of the original

triangle (top) is seen from three fixation points, each in the neighborhood of a particular vertex. These views are indicated by the corresponding boundary-angle functions. For each fixation point, only the closest vertex and its neighborhood are detailed, while the other vertices are approximated roughly. The reconstructed boundary-angle function (bottom) consists of the "best" contribution from each space-variant view.
B: a silhouette of an airplane, viewed from three fixation points, selected (by hand) because they are near areas containing many details. Details as in **A**.

Figure 4. A: Images (left) and the corresponding boundary-angle functions (right). The top row shows the original image and function; the next three rows represent three fixation points (denoted by small stars on the images), and the bottom row shows the integrated image and function. The fixation points, which are selected automatically, are the spatial locations that correspond to the three largest values of the original boundary-angle function (denoted by bars under the function).
B: Results of the modified algorithm. The fixation points are chosen by the maximum of $\frac{\Theta(U_i)}{R(U_i)}$.

Figure 5. Converging rate of the algorithm, as depicted by the difference Δ_n between successive blended figures. Left: blended figures after 1,2,3..8 fixation points. Right: Δ_n versus number of fixation points. Δ_n is the mean square error between two succesive figures, and is normalized to [0,1].

Figure 6. Figure 6A shows a wide angle fish eye view of a scene in the hall of our laboratory. A ladder is to the right, an eye chart is in the very center of the frame (almost invisible). The original version of this scene was digitized to an effective resolution of 16000x16000 pixels by a polar coordinate mosaic technique. A "blow-up" of the central region of this original frame is shown in figure 6B. This is an eye-chart, and the distance to the chart was twenty-feet. In the original, line 7 of the chart could be easily read, indicated an effective "acuity" of 20/30, or about 1.5 minutes of arc.

The purpose of this work was to simulate a wide angle scene (about 100 degrees), roughly comparable to human vision, at human visual acuity. Figure 6C shows this scene, blurred by a space variant filter which is modeled after human visual acuity. Figure 6D shows the image of 6A, modeled in terms of a complex logarithmic model[7] of human visual cortex. The eye-chart occupies almost half of the surface of visual cortex, although it occupies a tiny fraction of the original scene. The ladder, and the windows of the original are compressed to almost the same size as the centrally fixated letters of the eye-chart. This illustrates the tremendous space variant compression of human vision. Variations in linear size of about 100^2:1 (10^4 in solid angle) occur from the center to the periphery of the human visual system.

References

[1] B.M. Dow, A.Z. Snyder, R.G. Vautin, and R. Bauer, "Magnification factor and receptive field size in foveal striate cortex of monkey," *Exp. Brain Res.*, vol. 44, pp. 213-228, 1981.

[2] Martin Levine, *Vision in Man and Machine*, McGraw-Hill, New York, 1985.

[3] D. Noton and L. Stark, "Scanpaths in saccadic eye movements while viewing and recognizing patterns," *Vision Research*, vol. 11, pp. 929-942, 1971.

[4] Rayner, K., "Eye movement in reading and information processing," *Psychological Bulletin*, vol. 85, pp. 618-660, 1978.

[5] C.W. Richard and H. Hemami, "Identification of 3D objects using Fourier descriptors of the boundary curve," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. SMC-4 No. 4, pp. 371-378, 1974.

[6] E.L. Schwartz, "On the mathematical structure of the retinotopic mapping of primate striate cortex," *Science*, vol. 227, p. 1066, 1985.

[7] E.L. Schwartz, "Computational anatomy and functional architecture of striate cortex: a spatial-mapping approach to perceptual coding," *Vision Research*, vol. 20, pp. 645-670, 1980.

[8] E.L. Schwartz, R. Desimone, T. Albright, and C.G. Gross, "Shape recognition and inferior temporal neurons," *Proceedings of the National Academy of Sciences*, 1983.

[9] R.W. Sperry, *J. Comp. Physiology*, vol. 43, pp. 482-489, 1950.

[10] R.B. Tootel, M.S. Silverman, E. Switkes, and R. deValois, "Deoxyglucose, retinotopic mapping and the complex log model in striate cortex," *Science*, vol. 227, p. 1066, 1985.

[11] D.C. VanEssen, W.T. Newsome, and J.H.R. Maunsell, "The visual representation in striate cortex of the macaque monkey: Assymetries, anisotropies, and individual variability.," *Vision Research*, vol. 24, pp. 429-448, 1984.

[12] E. Wolfson and E.L. Schwartz, "Space-variant image-processing II: A truncated-pyramid algorithm for giga-pixel image-warping," *Comp.Neuro.Tech.Rep.*, vol. CNS-TR-23-86, NYU Med. Ctr./Computational Neuroscience Laboratories, 1986.

[13] E. Wolfson, Y. Yeshurun, and E.L. Schwartz, "Space-variant image-processing II: Image-blending of multi-fixation logarithmic views," *Comp.Neuro.Tech.Rep.*, vol. CNS-TR-10-86, NYU Med. Ctr./Computational Neuroscience Laboratories, 1986.

Figure 1a

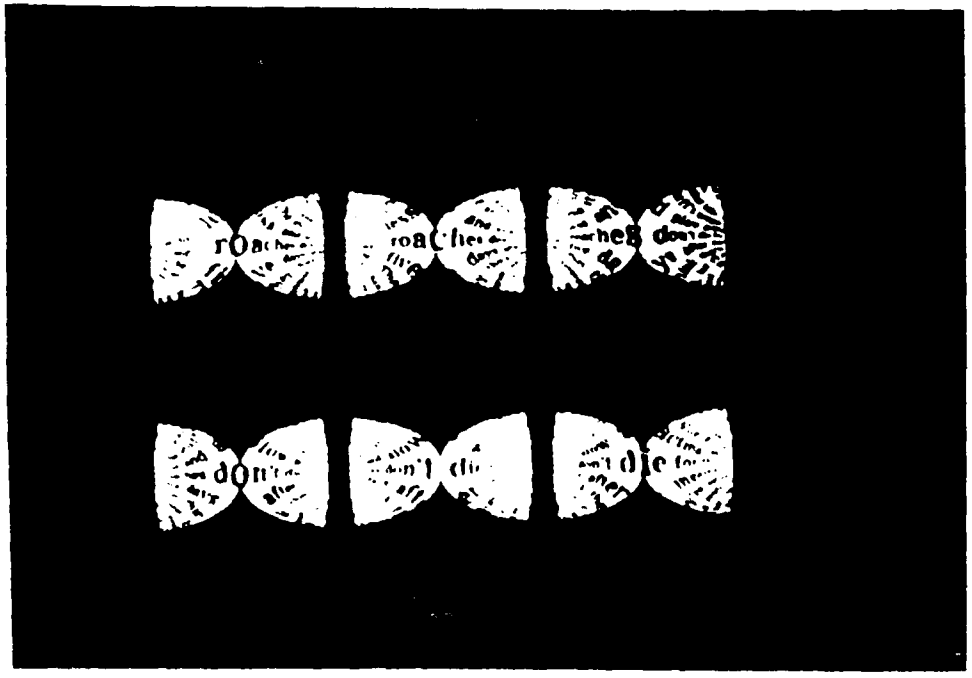


Figure 1b

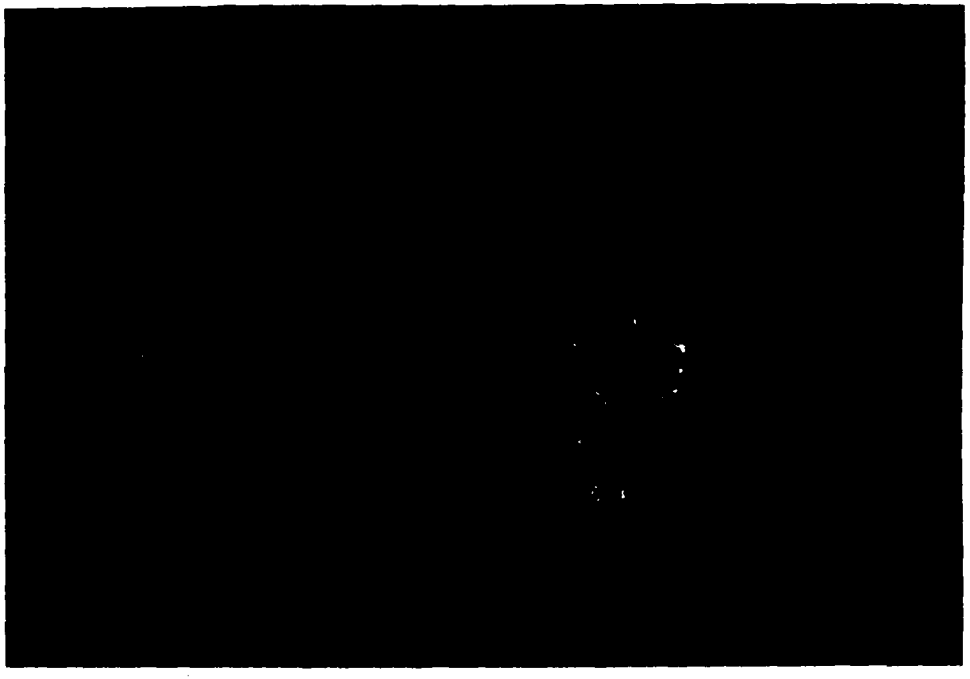


FIGURE 2a

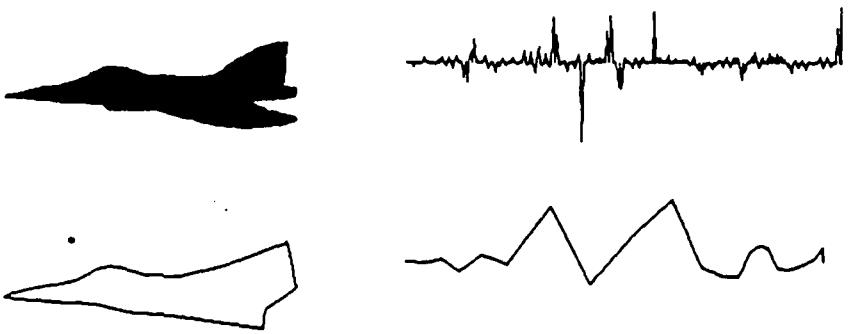


FIGURE 2b

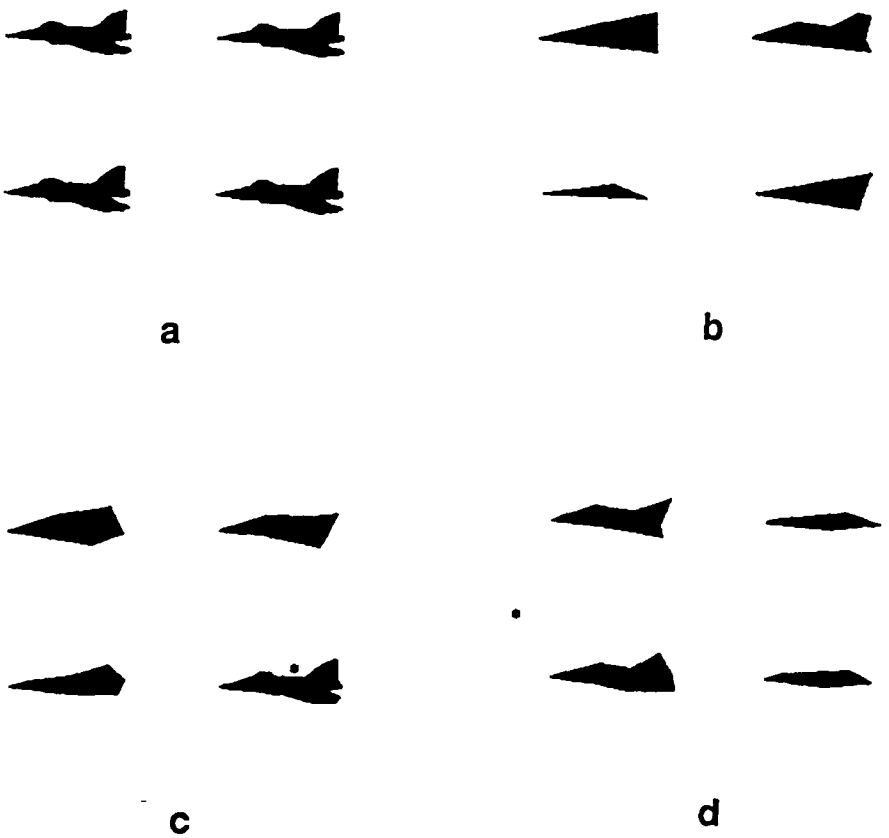


FIGURE 3a

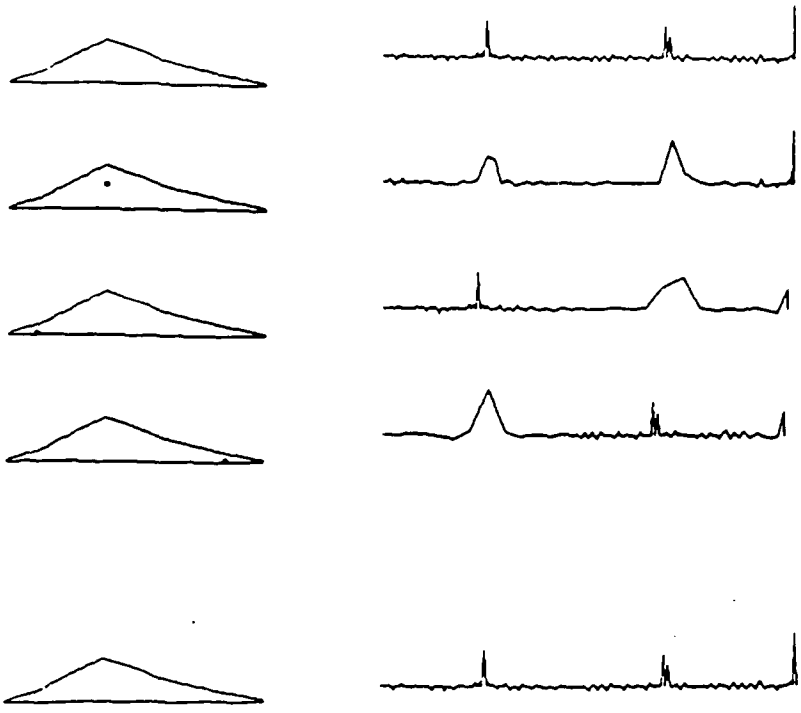


FIGURE 3b

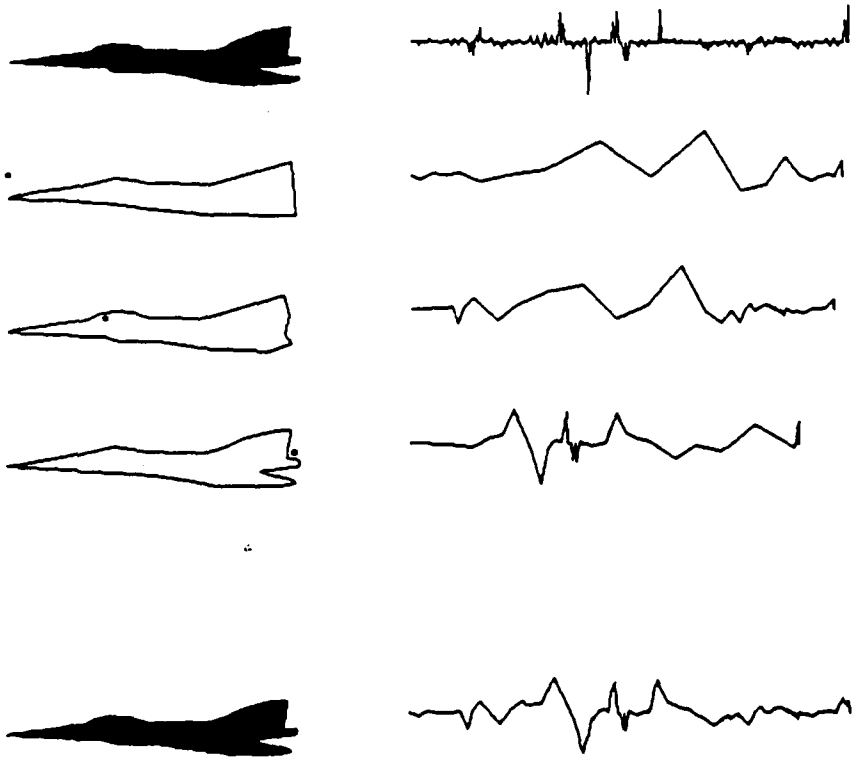


FIGURE 4a

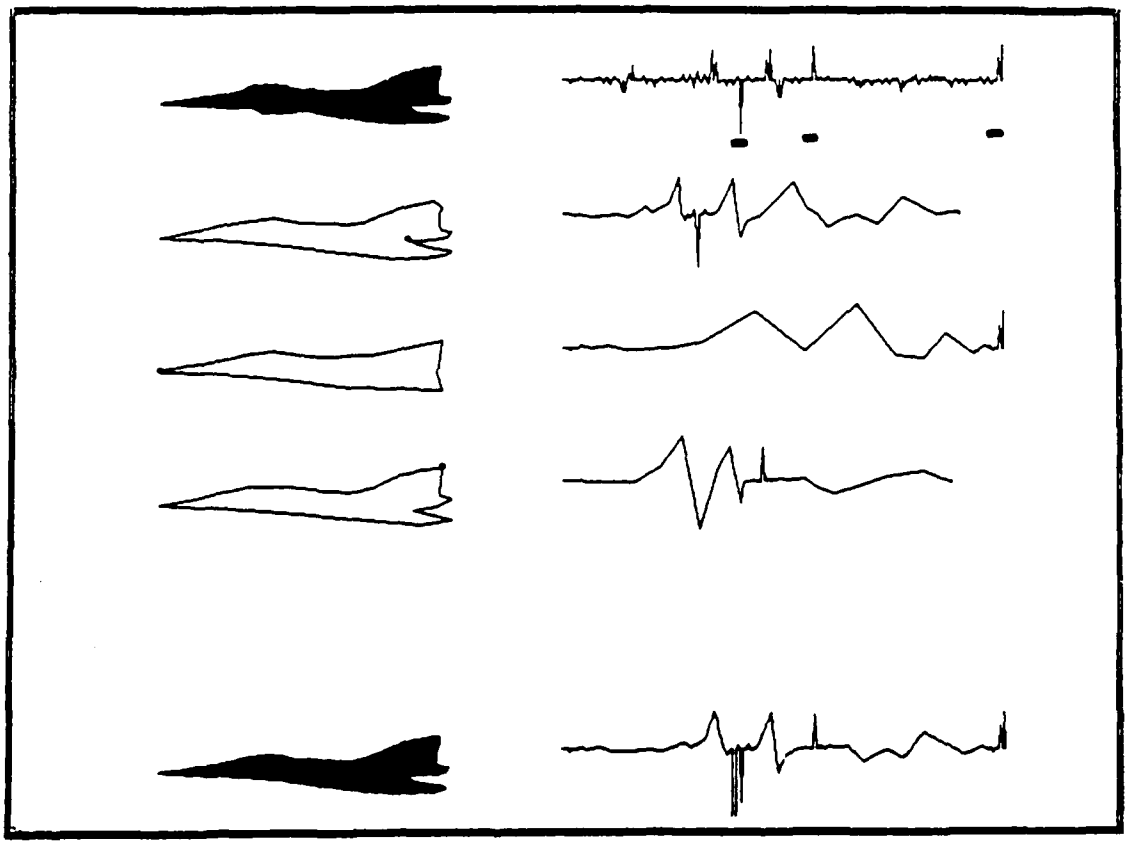
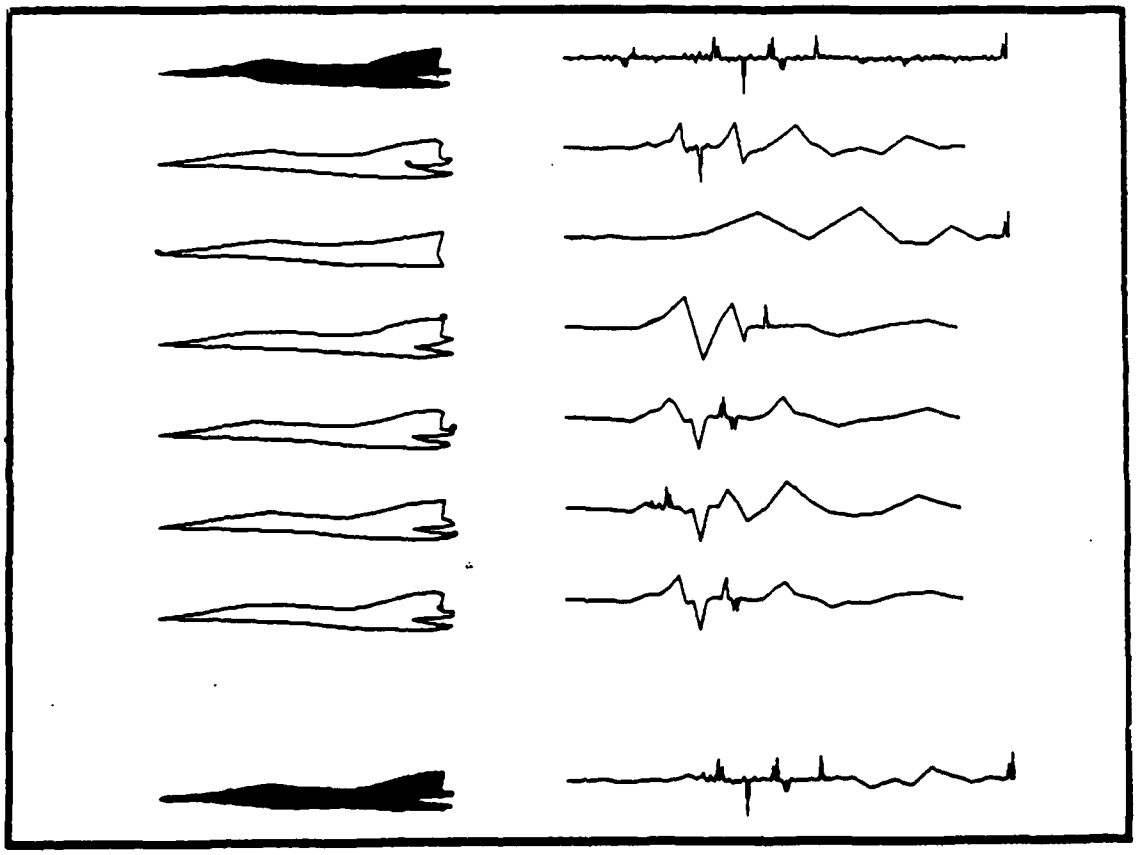
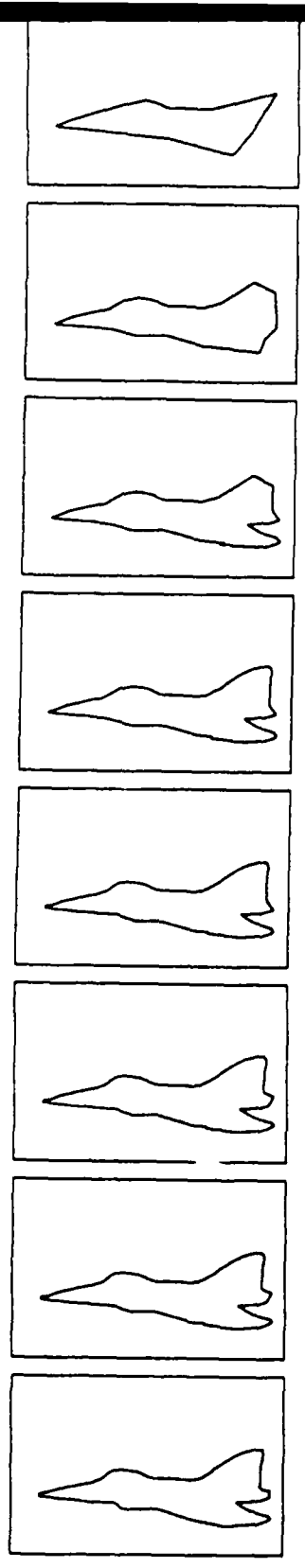


FIGURE 4b





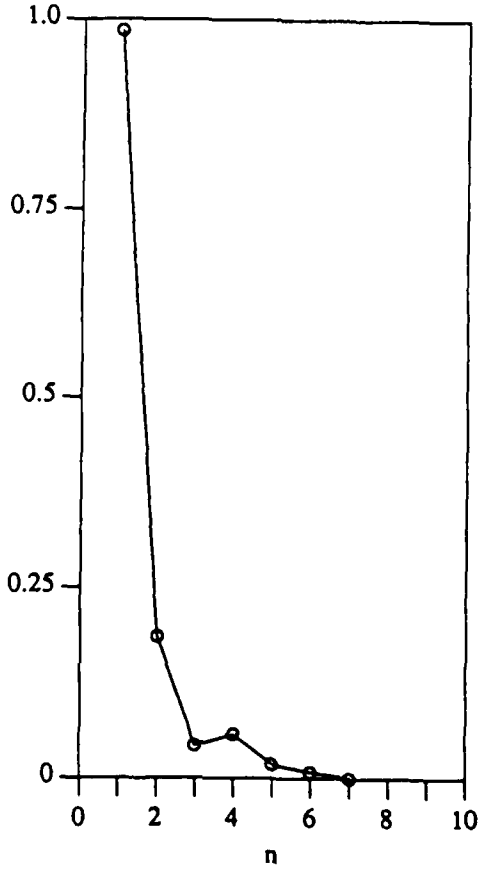
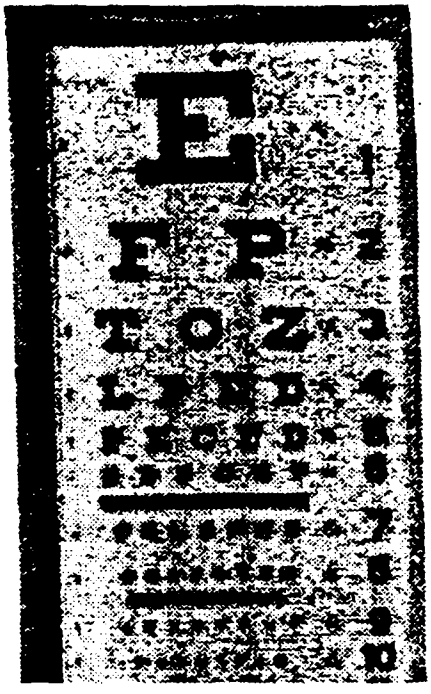


FIGURE 5

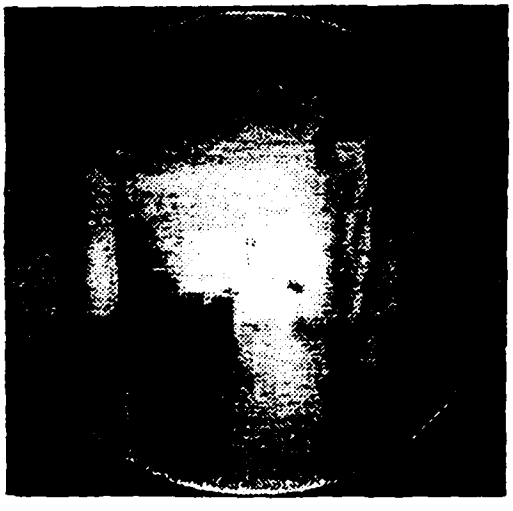
Figure 6

A

B



C



D